REVIEW

1. What is a linkage group? What is its relationship to a chromosome? How can one determine the number of linkage groups in a species?
2. How do genetic mutations help in mapping the location of genes on a chromosome?
3. What is meant by incomplete linkage? What does this have to do with pairing of homologous chromosomes during meiosis?
4. How do polytene chromosomes of an insect differ from normal chromosomes?

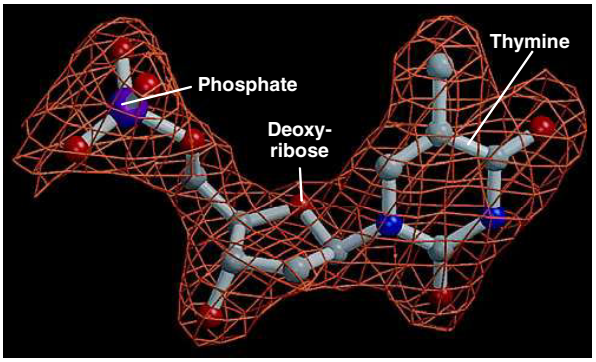
10.3 | The Chemical Nature of the Gene

Classical geneticists uncovered the rules governing the transmission of genetic characteristics and the relationship between genes and chromosomes. In his acceptance speech for the Nobel Prize in 1934, T. H. Morgan stated, "At the level at which the genetic experiments lie, it does not make the slightest difference whether the gene is a hypothetical unit or whether the gene is a material particle." By the 1940s, however, a new set of questions was being considered, foremost of which was, "What is the chemical nature of the gene?" The experiments answering this question are outlined in the Experimental Pathways at the end of this chapter. Once it became evident that genes are made of DNA, biologists were faced with a host of new questions. These questions will occupy us for the remainder of the chapter.

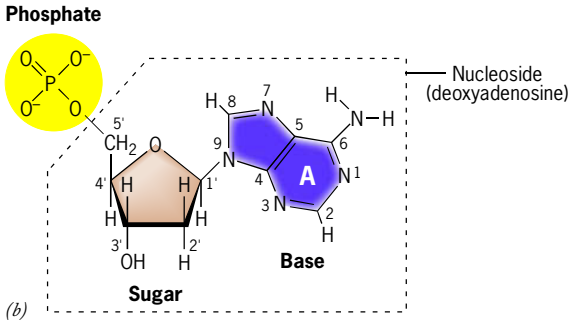
The Structure of DNA

To understand the workings of a complex macromolecule—whether it is a protein, polysaccharide, lipid, or nucleic acid—it is essential to know how that molecule is constructed. The mystery of DNA structure was investigated by a number of laboratories in both the United States and England in the early 1950s and was solved by James Watson and Francis Crick at Cambridge University in 1953. Before describing their proposed DNA structure, let us consider the facts available at the time.

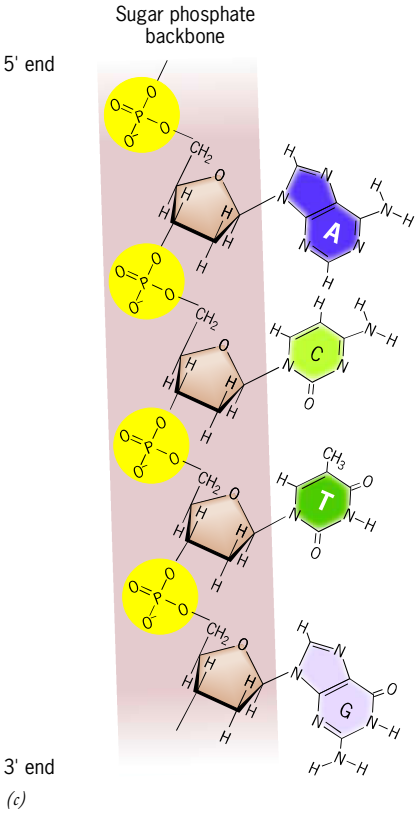
Base Composition The basic building block of DNA was known to be a **nucleotide** (Figure 10.9*a*,*b*) consisting of the five-carbon sugar *deoxyribose* to which one phosphate is esterified at



(a)



(b)



(c)

Figure 10.9 **The chemical structure of DNA.** (*a*) Model of a DNA nucleotide containing the base thymine; the molecule is deoxythymidine 5′-monophosphate (dTMP). The netlike cage represents the electron density of the atoms that make up the molecule. (*b*) Chemical structure of a DNA nucleotide containing the base adenosine; the molecule is deoxyadenosine 5′-monophosphate (dAMP). A nucleotide is composed of a nucleoside linked to a phosphate; the nucleoside portion of the molecule (i.e., deoxyadenosine) is enclosed by the dashed line. (*c*) The chemical structure of a small segment of a single DNA strand showing all four nucleotides. (A: FROM ARNON LAVIE ET AL., NATURE STR. BIOL. 4:604, 1997, FIG. 2A. REPRINTED BY PERMISSION OF MACMILLAN PUBLISHERS LIMITED.)

the 5′ position of the sugar ring and one nitrogenous base is attached at the 1′ site.[2] Two types of nitrogenous bases are present in a nucleic acid: **pyrimidines**, which contain a single ring, and **purines**, which contain two rings (Figure 10.9*c*). DNA contains two different pyrimidines, *thymine* (*T*) and *cytosine* (*C*), and two different purines, *guanine* (*G*) and *adenine* (*A*). The nucleotides are covalently linked to one another to form a linear polymer, or *strand*, with a backbone composed of alternating sugar and phosphate groups joined by *3′-5′-phosphodiester bonds* (Figure 10.9*c*). The bases attached to each sugar were thought to project from the backbone like a column of stacked shelves.

A nucleotide has a polarized structure: one end, where the phosphate is located, is called the *5′ end* (pronounced "five-prime end"), while the other end is the *3′ end* (Figure 10.9*b*). Because each of the stacked nucleotides in a strand faces the same direction, the entire strand has a direction. One end of the strand is the 3′ end, the other is the 5′ end (Figure 10.9*c*). X-ray diffraction analysis indicated that the distance between the nucleotides of the stack was 3.4 Å (0.34 nm) and suggested the presence of a large structural repeat every 3.4 nm.

As discussed on page 421, DNA was thought for many years to consist of a simple repeating tetranucleotide (e.g., —ATGCATGCATGC—), which made it unlikely to serve as an informational macromolecule. In 1950, Erwin Chargaff of Columbia University reported an important finding that delivered the final blow to the tetranucleotide theory and provided vital information about DNA structure. Chargaff, believing that the sequence of nucleotides of the DNA molecule held the key to its importance, determined the relative amounts of each base in various samples of DNA, that is, the **base composition** of the samples. Base composition analysis of a DNA sample was performed by hydrolyzing the bases from their attached sugars, separating the bases in the hydrolysate by paper chromatography, and determining the amount of material in each of the four spots to which the bases migrated.

If the tetranucleotide theory were correct, each of the four bases in a DNA sample would be present as 25 percent of the total number. Chargaff found that the ratios of the four component bases were quite variable from one type of organism to another, often being very different from the 1:1:1:1 ratio predicted by the tetranucleotide theory. For example, the A:G ratio of the DNA of a tubercle bacillus was 0.4, whereas the A:G ratio of human DNA was 1.56. It made no difference which plant or animal tissue was used as the source of the DNA; the base composition remained constant for that species. Amid this great variability in base composition of different species' DNA, an important numerical relationship was discovered. The number of purines always equaled the number of pyrimidines in a given sample of DNA. More specifically, the number of adenines always equaled the number of thymines, and the number of guanines always equaled the number of cytosines. In other words, Chargaff discovered the following rules of DNA base composition:

$$[A] = [T], [G] = [C], [A] + [T] \neq [G] + [C]$$

Chargaff's findings put the DNA molecule in a new light, giving it specificity and individuality from one organism to another. The significance of the base equivalencies, however, remained obscure.

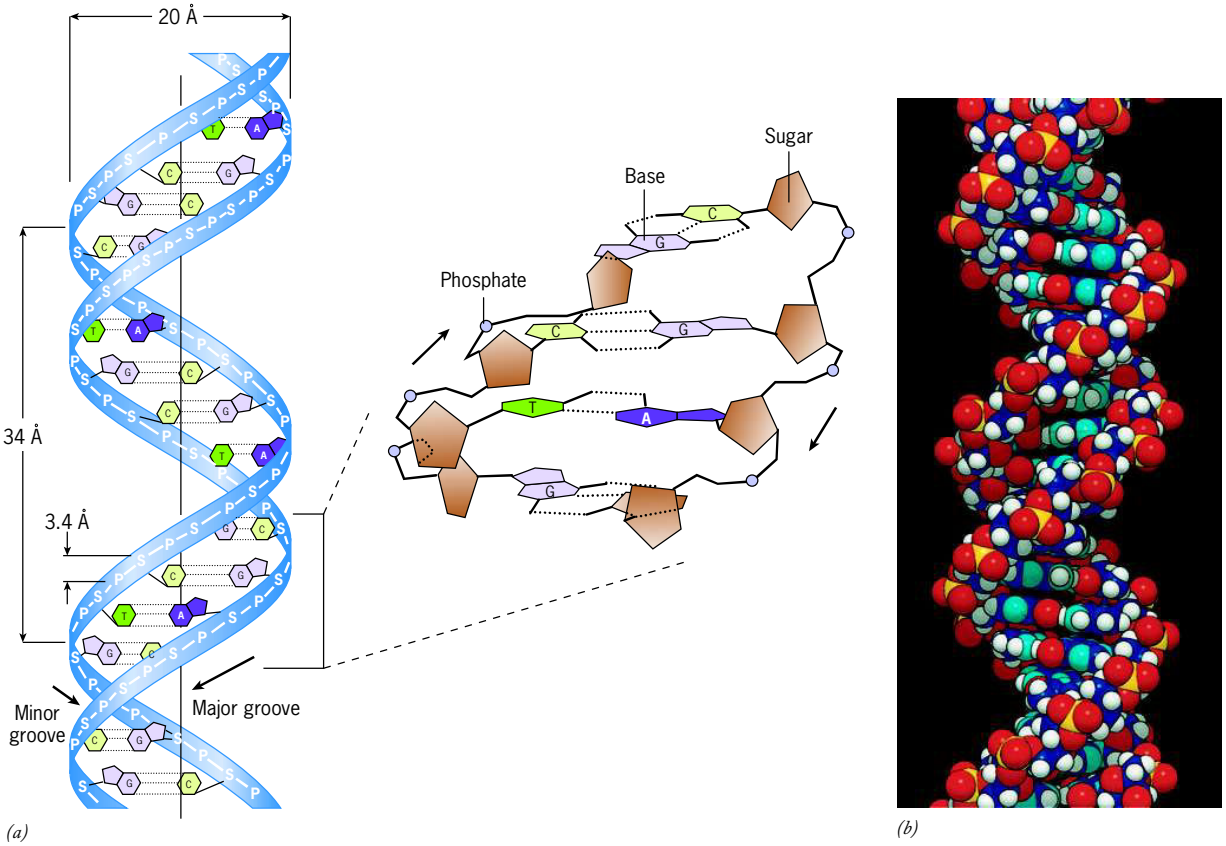
The Watson-Crick Proposal

When protein structure was discussed in Chapter 2, the importance of secondary and tertiary structure as determinants of the protein's activity was stressed. Similarly, information about the three-dimensional organization of DNA was needed if its biological activity was to be understood. Using X-ray diffraction data (obtained by Rosalind Franklin and Maurice Wilkins at King's College London) and models constructed from cutouts of the four types of nucleotides, Watson and Crick proposed a structure of DNA that included the following elements (Figure 10.10):

1. The molecule is composed of two chains of nucleotides. This conclusion followed on the heels of an erroneous proposal by Linus Pauling, who had suggested that DNA was composed of three nucleotide strands.
2. The two chains spiral around each other to form a pair of right-handed helices. In a right-handed helix, an observer looking down the central axis of the molecule would see that each strand follows a clockwise path, as it moves away from the observer. The helical nature of DNA was revealed in the pattern of spots produced by Franklin's X-ray diffraction image (seen on page 386), which was shown to Watson during a visit to King's College.

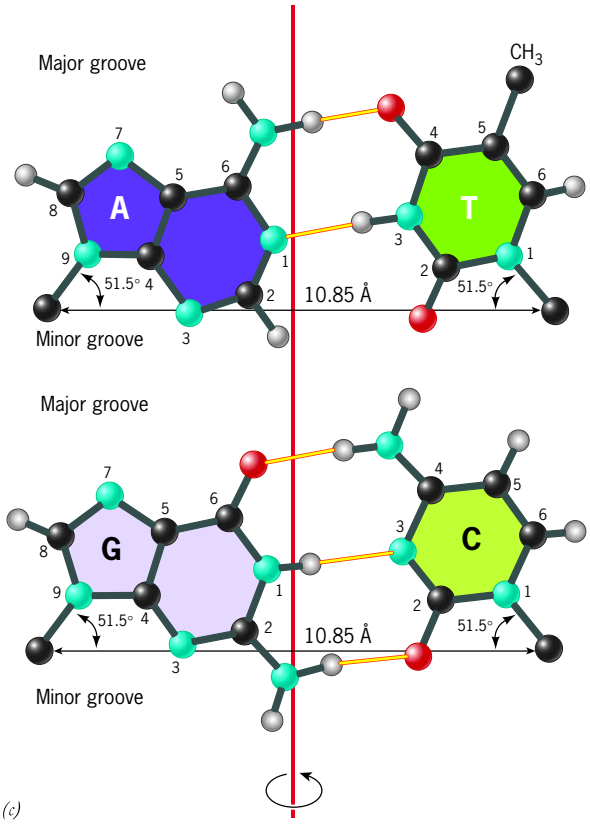
Figure 10.10 **The double helix. (*opposite*)** (*a*) Schematic representation of the DNA double helix. (*b*) Space-filling model of the B form of DNA. (*c*) The Watson-Crick base pairs. The original model showed both A-T and G-C pairs with two hydrogen bonds; the third hydrogen bond in the G-C pair was subsequently identified by Linus Pauling. (*d*) Electron micrograph of DNA being released from the head of a T2 bacteriophage. This linear DNA molecule (note the two free ends) measures 68 μm in length, approximately 60 times longer than the phage head in which it is contained. (B: COURTESY OF NELSON MAX, LAWRENCE LIVERMORE NATIONAL LABORATORY AND THE DEPARTMENT OF ENERGY. C: FIGURE 28.6 ON PAGE 854 FROM VOET & VOET, BIOCHEMISTRY, 2E; COPYRIGHT 1995, JOHN WILEY & SONS, INC. THIS MATERIAL IS REPRODUCED WITH PERMISSION OF JOHN WILEY & SONS, INC. D: FROM A. K. KLEINSCHMIDT, ET AL., BIOCHIM. BIOPHYS. ACTA 61:861, 1962, WITH PERMISSION FROM ELSEVIER.)

[2]It is useful to introduce a bit of terminology at this point. A molecule consisting simply of one of the four nitrogenous bases of Figure 10.9 linked to a pentose sugar moiety is known as a nucleoside. If the sugar is deoxyribose, the nucleoside is a deoxyribonucleoside. There are four major deoxyribonucleosides distinguished by the attached base: deoxyadenosine, deoxyguanosine, deoxythymidine, and deoxycytidine. If the nucleoside has one or more attached phosphate groups (generally at the 5′ position, but alternatively at the 3′ position), the molecule is a nucleotide. There are nucleoside 5′-monophosphates, nucleoside 5′-diphosphates, and nucleoside 5′-triphosphates, depending on the number of phosphates in the molecule. Examples of each are deoxyadenosine 5′-monophosphate (dAMP), deoxyguanosine 5′-diphosphate (dGDP), and deoxycytidine 5′-triphosphate (dCTP). A similar set of nucleosides and nucleotides involved in RNA metabolism contain the sugar ribose rather than deoxyribose. The nucleotides employed in energy metabolism, such as adenosine triphosphate (ATP), are ribose-containing molecules.

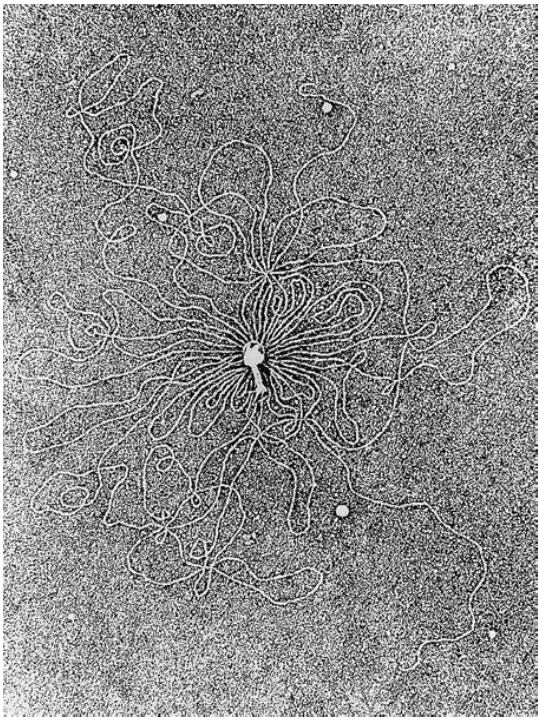

20 Å
34 Å
3.4 Å
Minor groove
Major groove
Sugar
Base
Phosphate

(a)

(b)


Major groove
CH3
A
T
51.5°
10.85 Å
Minor groove
Major groove
G
C
51.5°
10.85 Å
Minor groove

(c)



(d)

3. The two chains comprising one double helix run in opposite directions; that is, they are *antiparallel*. Thus, if one chain is aligned in the 5′ → 3′ direction, its partner must be aligned in the 3′ → 5′ direction.
4. The –sugar–phosphate–sugar–phosphate– backbone of each strand is located on the outside of the molecule with the two sets of bases projecting toward the center. The phosphate groups give the molecule a large negative charge.
5. The bases occupy planes that are approximately perpendicular to the long axis of the molecule and are, therefore, stacked one on top of another like a pile of plates. Hydrophobic interactions and van der Waals forces (page 36) between the stacked, planar bases provide stability for the entire DNA molecule. Together, the helical turns and planar base pairs cause the molecule to resemble a spiral staircase. This manner of construction is evident in the chapter-opening photograph, which shows the original Watson-Crick model.
6. The two strands are held together by hydrogen bonds between each base of one strand and an associated base on the other strand. Because individual hydrogen bonds are weak and easily broken, the DNA strands can become separated during various activities. But the strengths of hydrogen bonds are additive, and the large numbers of hydrogen bonds holding the strands together make the double helix a stable structure.
7. The distance from the phosphorus atom of the backbone to the center of the axis is 1 nm (thus the width of the double helix is 2 nm).
8. A pyrimidine in one chain is always paired with a purine in the other chain. This arrangement produces a molecule that is 2 nm wide along its entire length.
9. The nitrogen atoms linked to carbon 4 of cytosine and carbon 6 of adenine are predominantly in the amino (NH_2) configuration (Figure 10.9*c*) rather than the imino (NH) form. Similarly, the oxygen atoms linked to carbon 6 of guanine and carbon 4 of thymine are predominantly in a keto (C═O) configuration rather than the enol (C—OH) form. These structural restrictions on the configurations of the bases suggested that adenine was the only purine structurally capable of bonding to thymine and that guanine was the only purine capable of bonding to cytosine. Therefore, the only possible pairs were A-T and G-C (Figure 10.10*c*). This fit perfectly with the base composition analysis carried out by Chargaff whose results were communicated to Watson and Crick during a meeting of the three scientists in Cambridge in 1952. Because an A-T and G-C base pair had the same geometry (Figure 10.10*c*), there were no restrictions on the sequence of bases; a DNA molecule could have any one of an unlimited variety of nucleotide sequences.
10. The spaces between adjacent turns of the helix form two grooves of different width—a wider *major groove* and a more narrow *minor groove*—that spiral around the outer surface of the double helix. Proteins that bind to DNA often contain domains that fit into these grooves. In many cases, a protein bound in a groove is able to read the sequence of nucleotides along the DNA without having to separate the strands.
11. The double helix makes one complete turn every 10 residues (3.4 nm), or 150 turns per million daltons of molecular mass.
12. Because an A on one strand is always bonded to a T on the other strand, and a G is always bonded to a C, the nucleotide sequences of the two strands are always fixed relative to one another. Because of this relationship, the two chains of the double helix are said to be **complementary** to one another. For example, A is complementary to T, 5′-AGC-3′ is complementary to 3′-TCG-5′, and one entire chain is complementary to the other. As we shall see, complementarity is of overriding importance in nearly all the activities and mechanisms in which nucleic acids are involved.

The Importance of the Watson-Crick Proposal From the time biologists first considered DNA as the genetic material, it was expected to fulfill three primary functions (Figure 10.11):

1. ***Storage of genetic information.*** As genetic material, DNA must contain a stored record of instructions that determine all the inheritable characteristics that an organism exhibits. In molecular terms, DNA must contain the information for the specific order of amino acids in all the proteins that are synthesized by an organism.
2. ***Replication and inheritance.*** DNA must contain the information for synthesis of new DNA strands (replication). DNA replication allows genetic instructions to be transmitted from one cell to its daughter cells and from one individual to its offspring.
3. ***Expression of the genetic message.*** DNA is more than a storage center; it is also a director of cellular activity. Consequently, the information encoded in DNA has to be expressed in some form that can take part in events that are taking place within the cell. More specifically, the information in DNA must be used to direct the order by which specific amino acids are incorporated into a polypeptide chain.

The Watson-Crick model of DNA structure suggested a way in which the first two of these three genetic functions might be accomplished. The model strongly supported the suspicion that the information content of DNA resided in the linear sequence of its bases. A given segment of DNA would correspond to each gene. The specific sequence of nucleotides in that segment would dictate the sequence of amino acids in a corresponding polypeptide. A change in the linear sequence of nucleotides within that segment would correspond to an inheritable mutation in that gene. Differences in nucleotide sequence were presumed to form the basis of genetic variation, whether between individuals of a species or between species themselves. As discussed in the next chapter, another decade would pass before biologists would learn the mechanism by which a sequence of amino acids is encoded by a sequence of DNA nucleotides.

As for the second function, Watson and Crick's initial publication of DNA structure included a proposal as to how such a molecule might replicate. This may have been the first time that

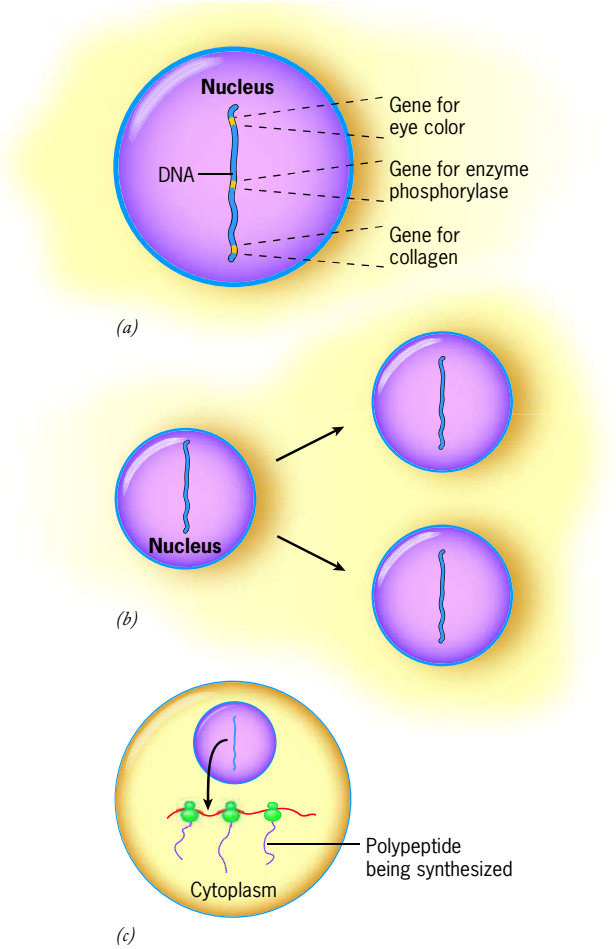


Figure 10.11 **Three functions required of the genetic material.** (*a*) DNA must contain the information that encodes inheritable traits. (*b*) DNA must contain the information that directs its own duplication. (*c*) DNA must contain the information that directs the assembly of specific proteins.

a study of molecular structure led directly to a hypothesis of a basic molecular mechanism. Watson and Crick proposed that during replication, the hydrogen bonds holding the two strands of the DNA helix were sequentially broken, causing the gradual separation of the strands much like the separation of two halves of a zipper. Each of the separated strands, with its nitrogenous bases exposed, would then serve as a *template* directing the order in which complementary nucleotides become assembled to form the complementary strand. When complete, the process would generate two double-stranded DNA molecules that were (1) identical to one another and (2) identical to the original DNA molecule. According to the Watson-Crick proposal, each DNA double helix would contain one strand from the original DNA molecule and one newly synthesized strand (see Figure 13.1). As we will see in Chapter 13, the Watson-Crick proposal fared quite well in predicting the mechanism of DNA replication. Of the three primary functions listed above, only the mechanism by which DNA governs the assembly of a specific protein remained a total mystery.

Not only was the elucidation of DNA structure important in its own right, it provided the stimulus for investigating all the activities in which the genetic material must take part. Once the model for its structure was accepted, any theory of a genetic code, DNA synthesis, or information transfer had to be consistent with that structure.

DNA Supercoiling

In 1963, Jerome Vinograd and his colleagues at the California Institute of Technology discovered that two closed, circular DNA molecules of identical molecular mass could exhibit very different rates of sedimentation during centrifugation (Section 18.9). Further analysis indicated that the DNA molecule sedimenting more rapidly had a more compact shape because the molecule was twisted upon itself (Figure 10.12*a*,*b*), much like a rubber band in which the two ends are twisted in opposite directions or a tangled telephone cord after extended use. DNA in this state is said to be **supercoiled**. Because supercoiled DNA is more compact than its relaxed counterpart, it occupies a smaller volume and moves more rapidly in response to a centrifugal force or an electric field (Figure 10.12*c*).

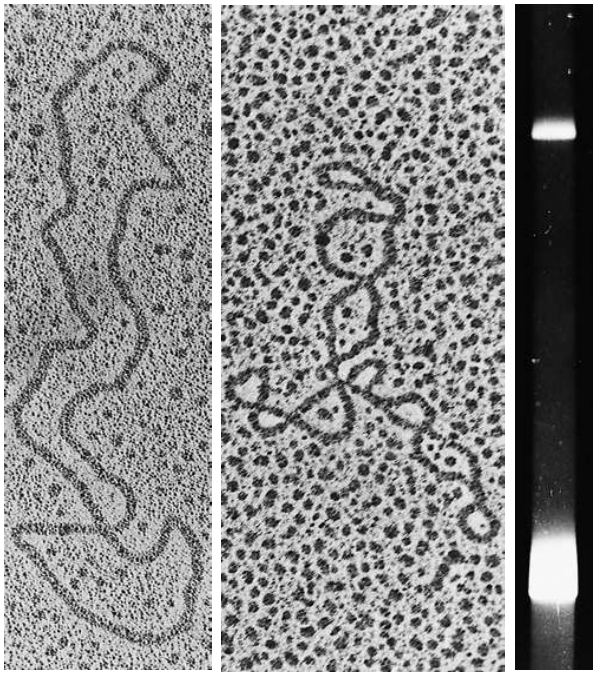


Figure 10.12 **Supercoiled DNA.** (*a*,*b*) Electron micrographs showing the differences in conformation between a relaxed, circular molecule of phage DNA (*a*) and the same type of molecule in a supercoiled state (*b*). (*c*) When a mixture of relaxed and supercoiled SV40 DNA molecules are subjected to gel electrophoresis, the highly compact, supercoiled form (seen at the bottom of the gel) moves much more rapidly than the relaxed form. The DNA molecules are visualized by staining the gel with ethidium bromide, a fluorescent molecule that inserts itself into the double helix. (A,B: COURTESY OF JAMES C. WANG; C: FROM WALTER KELLER, PROC. NAT'L. ACAD. SCI. U.S.A. 72:2553, 1975.)